

Dialect Geography of Khmer in Northeastern Thailand: An Analysis of Regional Variation and Sub-dialect Classification¹

Ponchanok Saenthaweesuk Seedabut^{a*} and Pattama Patpong^a

^a*Research Institute for Languages and Cultures of Asia, Mahidol University,
Nakhon Pathom, Thailand*

^{*}*Corresponding Author. Email: ponchanok.sdb@gmail.com*

Received: February 12, 2024

Revised: June 5, 2024

Accepted: September 10, 2024

Abstract

This study investigates the dialect geography of the Khmer dialect in northeastern Thailand, employing Chambers and Trudgill's dialectology approach and applying GIS for linguistic map analysis. This study is a synchronic analysis of Khmer phonological variation which aims to identify sub-dialects spoken in Thailand's northeastern provinces, focusing on regional variations in initials, vowels, and finals. The findings result in a classification of Khmer into two main groups: a Western Khmer and an Eastern Khmer group, including fifteen sub-dialects within the studied region concerning the overlap on linguistic maps of the initial consonant, vowel, and final consonant.

Keywords: Khmer, sub-dialect, regional variation, dialect, dialect geography

Introduction

Khmer, an Austroasiatic language within the Khmeric branch, is the official language of Cambodia. While Khmer dialects extend to parts of Thailand, Laos, and Vietnam, this study focuses on those spoken in

¹ This article is part of a Ph.D. dissertation at Mahidol University, titled "Dialect Geography of Khmer Phonology in the Northeast of Thailand." This research received a specific grant from the Royal Golden Jubilee (RGJ) Ph.D. Programme (Grant No. PHD/0166/2560) through the National Research Council of Thailand (NRCT), the Thailand Research Fund (TRF), and the Thailand Science Research and Innovation (TSRI).

northeastern Thailand. Smalley (1964, 1976 cited in Premsrirat, 1997: 129) introduced the term “Northern Khmer” to refer to the Khmer language spoken in Thailand. In 1990, the northern Khmer people agreed to officially refer to themselves and their language as “Thailand Khmer” (Premsrirat and Srichampa, 1990).

According to the Ethnolinguistic Maps of Thailand’s latest survey, around 1.4 million native Khmer speakers exist in Thailand (Premsrirat et al., 2004: 84). They are dispersed across the northeast, east, and central regions, encompassing provinces such as Surin, Buri Ram, Si Sa Ket, Ubon Ratchathani, Nakhon Ratchasima, Maha Sarakham, Roi Et, Prachin Buri, Trat, Chanthaburi, Chachoengsao, and Sa Kaeo (Premsrirat and Phuithsakrit, 1996: 15). Most Khmer dialect speakers reside in northeast Thailand, notably in provinces like Surin, Buri Ram, and Si Sa Ket, which constitute 47, 27.6, and 26.2 percent, respectively, of the entire Khmer population in Thailand (Vail, 2007: 111; Hattaway, 2004: 133). These communities have established roots in the region over several generations. In other parts of the country, smaller Khmer-speaking communities, with fewer than 10,000 members, are found, particularly in the eastern and central regions (Premsrirat and Phuithsakrit, 1996). Noteworthy Khmer-speaking populations exist in central provinces like Ratchaburi (Kraisornrat, 2012: 22; Treesawat, 2012: 26-28), Suphan Buri (Suphan Buri National Archives Commemoration, 2018), and Nakhon Pathom (Cultural Center of Mueang Nakhon Pathom, 2013). Older speakers are more proficient in Khmer in these areas, while the younger generation tends to converse in Thai (Cultural Center of Mueang Nakhon Pathom, 2013). The Khmer-speaking history in the eastern part of the country reveals that communities in Chanthaburi and Sa Kaeo provinces have resided there for over three generations. In Chachoengsao province, speakers’ ancestors migrated to Thailand around 1781-1782 during the Cambodian war invasion by the King of Thonburi (Premsrirat et al., 2004: 157). Along with long-established communities, recent migrants and immigrant workers, drawn by employment opportunities, are dispersed throughout these regions (Southong, 2002: 2; Southong, 2013: 21).

Given the northeast’s dense Khmer-speaking population, the study focuses solely on Khmer varieties in this region for fieldwork.

Prior linguistic studies on Khmer in Thailand have been conducted in various sub-fields, encompassing phonology, morphology, syntax, sociolinguistics, and literacy. While some focus on specific dialectal variations, such as final consonants (Premsrirat and Phuithsakrit, 1996) and vowels (Southong, 2013), none have explored the phonological aspects or identified sub-dialects in the northeastern region of Thailand. Comparative overview studies are few. Instead, individual dialect descriptions have been conducted for Surin (Chantrupanth and Phromjakgarin, 1978; Phon-ngam, 1987; Premsrirat and Srichampa, 1990), Buri Ram (Yodmongkhon, 1986), Si Sa Ket (Seedabut, 2016), Chanthaburi (Southong, 2002), and Chachoengsao (Wichitkhachee, 1996). These studies show that Khmer dialects in Thailand (see Appendix for the summarized detail of each dialect’s phonemic systems and phoneme inventories) exhibit a range of 17 to 22 consonant phonemes, applicable as onsets, with only 11 to 14 phonemes permissible in word-final positions. By contrast, Chachoengsao and Si Sa Ket dialects feature 13 final consonant phonemes. Notably, a study on the Surin dialect by Chantrupanth and Phromjakgarin (1978) identified only 11 final consonant phonemes. The count of phonemes is contingent on syllable structure interpretation; for instance, Phon-ngam (1987) construed /ph/ as CC, where /p/ and /h/ are distinct phonemes, while Premsrirat and Srichampa (1990) considered /ph/ as an aspirated C. Moreover, the Chachoengsao dialect lacks the final /-r-/, while the Si Sa Ket dialect lacks the /-ʔ/.

Regarding vowels, Seedabut and Burusphat’s study (2017) highlights the Si Sa Ket dialect, which possesses the fewest vowel phonemes, while the Chanthaburi dialect has the highest number. Khmer dialects in Surin, Buri Ram, Chachoengsao, and Si Sa Ket maintain a clear vowel system, whereas the Chanthaburi dialect has two vowel systems, distinguishing between clear and creaky registers. Notably, each dialect’s vowel system tends to align with neighboring areas. Premsrirat (1997: 131) further describes the Khmer vowel system as

rather complex and atypical in the number of basic vowels and their positions, as evident in the abundance of vowels across these five dialects.

Phon-ngam (1987: 113-120) noted that most Khmer words consist of one or two syllables, categorized into monosyllabic and disyllabic structures. Monosyllabic words bear strong stress on their syllable. Khmer dialects exhibit three types of cluster consonant syllable structures: CC (e.g., /srɔʔ/ 'village'), CCC (e.g., /thnɔl/ 'road'), and CCCC (e.g., /pphle:c/ 'radiant'). Syllables are classified into major (always with primary stress), minor (always with weak stress), and presyllables (without stress). Khmer features sesquisyllabic structures (Thomas, 1992) in the presyllable, a unique trait in Austroasiatic languages. When consonants pose pronunciation challenges, the transitional vowel [ə] is inserted phonetically between them. During pronunciation, this sesquisyllabic has characteristics similar to disyllabic words but has the syllable structure of monosyllabic words (e.g., /pka:/ 'flower' pronounced as [ˈpəka:]). Khmer dialects have three intonation types: falling, rising, and rising-falling (Phon-ngam, 1987). Regarding register, Southong (2013) described Chanthaburi and Si Sa Ket Khmer dialects as exhibiting a clear and breathy contrast in their vowel systems.

Drawing from Seedabut's (2016) study, the Khmer dialect in Si Sa Ket province, situated in the northeastern part of Thailand, displays variations in finals (e.g., -r, -k, -c, -ʔ, -ŋ, -ɲ) and vowels. Notably, certain significant vowels (e.g., ɔ, ə, ʌ, ɔ, ɑ) exhibit free variation, aligning more closely with Standard Thai vowels. Standard Thai influences the Khmer population through the education system, government entities, and social media, impacting dialects. Within a community, dialects may differ, with neighboring villages showing distinct final consonant and vowel pronunciations. Additionally, speakers from nearby areas incorporate pitch as a distinctive feature in select words.

In studies examining Khmer's regional variation and dialectology, Premsrirat and Phuitsarakit (1996) explored Khmer's final variation with data from 37 sites, each representing a district, across eight provinces in Thailand. Southong's (2013) study on Khmer vowel

variation included data from 12 sites spanning 12 provinces. Both studies adopted the dialect geography framework, collecting data from representative communities in northeastern and eastern Thailand. The results, represented on paper maps, reveal vowel and final consonant phoneme variations. These studies, serving as methodological guides for the current research, expose gaps in previous Khmer studies, highlighting the need for identified sub-groups among Khmer speakers in Thailand.

Khmer dialects in Thailand are traditionally categorized by administrative divisions, primarily provinces and districts. Premsrirat (1997: 130) noted the common practice of naming dialects based on their geographical location, like Surin Khmer, Buri Ram Khmer, and Sangkha Khmer. Nonetheless, Seedabut's (2016) fieldwork observations reveal the existence of distinct sub-dialects within provinces, districts, and sub-districts.

Hence, to address existing gaps and provide updated insights into Khmer, the current study aims to investigate variations in Khmer spoken across northeastern provinces and subsequently identify sub-dialects in this region, mainly focusing on phonological aspects. The study adopts the theoretical framework of Chambers and Trudgill's (1980) dialectological approach. Dialectology explores linguistic patterns in language varieties or dialects, examining how phonological, morphological, or grammatical forms differ geographically within a language. Data is represented on linguistic maps, delineating dialect boundaries through bundles of isogloss, lines indicating linguistic feature distinctions between regions. Language variation, as defined by Crystal (1992), encompasses differences in grammar, pronunciation, or word selection influenced by factors such as region, gender, age, social class, education, occupation, and formality. These variations manifest across phonological, grammatical, and lexical levels.

Research Objective

To identify sub-dialects spoken in Thailand's northeastern provinces, focusing on regional variations in initials, vowels, and finals.

Research Methodology

Site Selection

This study's fieldwork spanned seven provinces in northeast Thailand: Surin, Buri Ram, Si Sa Ket, Ubon Ratchathani, Nakhon Ratchasima, Maha Sarakham, and Roi Et. Site selection focused on districts where Khmer speakers have resided for generations, with each sub-district chosen as a research site. The dialect in a sub-district indicates the entire community, and the village with the most Khmer speakers in each sub-district was selected as representative. Khmer speakers were identified across 202 communities in 34 districts within these seven provinces.

Participants

The dialectology approach for participant selection, as outlined by Chambers and Trudgill (1980: 33-35), involved choosing at least three native speakers from each sub-district based on specific criteria. This study collected data from a total of 618 participants using these guidelines:

Nonmobile: Participants must have been born and lived in the community from birth, with no extended periods outside the community exceeding one year.

Older: All participants had to be at least 35 years old, encompassing mainly elderly individuals from the grandparent's generation. However, middle-aged speakers were also included to broaden the age range and gather diverse insights into daily language use.

Rural: Participants needed to reside and work in rural areas, being native Khmer speakers using the dialect in their daily lives. Educational level and occupation were irrelevant.

Gender: Both male and female participants were required for a comprehensive study.

Additionally, articulators of participants needed to be in sufficiently good health to ensure clear and correct articulation, a crucial aspect of the selection process.

Data Collection

The data were collected from May to July 2019, and the collection process involved several key steps. First, we greeted the participants and inquired about their backgrounds to ensure they meet the study's criteria. Subsequently, we explained the study's purpose. If participants were willing and able to participate, the interview took approximately 30 minutes at a quiet and convenient venue. Second, data collection occurred with participants individually, using a digital voice recorder to document their responses based on a word list. The instruments for data collection were as follows.

1. Handheld GPS navigation device, map, and digital voice recorder

During fieldwork, the researcher used a GARMIN ETREX 10 handheld GPS navigation device alongside a Thailand highway map and navigation system for travel. Interviews were conducted using a SONY ICD-UX543F digital voice recorder.

2. Word list

A data collection tool, a basic word list of approximately 100 Khmer words, was employed. This selection was derived from previous studies on Khmer regional variation (Premsrirat and Phuissarakit, 1996; Southong, 2013) and phonology (Seedabut, 2016). Each word in the list consists of one or more variable phonemes. The presented word list is as follows:

3. GIS mapping tool

ArcGIS Online is a cloud-based geographic information system (GIS) platform developed by Esri. It facilitates the creation, sharing, and analysis of spatial data through a web-based interface. The platform supports a range of functionalities, including map creation, spatial analysis, data management, and collaborative tools for sharing geographic insights within and across organizations (Esri, n.d.). ArcGIS was employed to create linguistically accurate and detailed maps with multiple data layers, which facilitated the convenient display of Khmer distribution.

Table 1 Word list

| Khmer | English | Khmer | English | Khmer | English | Khmer | English |
|--------|----------------|----------|------------------------------|--------|--------------------------------|----------|-------------------|
| ci:k | to dig | naw | to stay, to live | we:ŋ | long | deŋ | to know |
| cci:k | ears | ʔaw | father, father- in-law | tiə | duck | deŋ | to chase |
| me:n | Yes | prəh | male | ciəh | clever | cu:n | to send |
| mi:w | cat | bo:h | abscess | ɲuər | to tremble | sne:c | bean |
| mel | salt | caŋ | want to | ntɔm | to stamp | phne:k | eyes |
| caŋkeh | waist | cməh | nose | miəm | finger | khmo:c | ghost |
| seh | horse | pdahsa:j | a cold | mamuət | a female medium | ce:k | banana |
| muət | mouth | lhəŋ | papaya | sək | hair | dac | alike, similar |
| cat | heart, mind | lɔ:k | monk | bok | to pound | tu:c | small |
| tpuəl | cheek | bɔ:n | four | tu:k | water | ʔac | stools |
| daj | arm, hand | khɔ:n | body | phək | to drink | hə:r | to fly |
| daj | soil | cɔ: | bad | tu:k | boat | khnɔr | jackfruit |
| baj | three | cu: | sour | cru:k | pig | hər | spicy |
| kbaj | buffalo | so:mtɔ:h | Sorry | cho:k | lotus | ka:r | work |
| truəj | fish | tɔ:h | penalty | sɔk | village | dər | to walk |
| sɔj | female | ʔa:rkun | Thank you | thliək | to fall | ju:r | long time |
| tanraj | elephant | ca: | to fry | tlɔ:k | coconut shell | chuər | to stand |
| pri: | forest | ɲuəm | dressed salad | luək | to sell | khme:r | Khmer |
| praj | salty | muən | chicken | tiŋ | to buy | ʔar | to saw |
| pra: | use | ruət | to run | pɲ | full | ci:r | fertilizer |
| chə: | wood | pruəŋ | hole | mi:ŋ | younger sister of mother | pi:r | two |
| chur: | hurt | puəŋ | egg | phli:ŋ | rain | pi:rtnap | twelve |
| kapɔ: | crocodile | puəh | snake | thmɲ | teeth | thwi:ər | door |
| kruəp | seed | puəh | stomach | hi:ŋ | bullfrog | wuər | vine |
| sɔaw | paddy | ptiəh | house | tneŋ | to ask | ɲno:ŋ | well (n.) |

Data Analysis

The collected data was analyzed by auditory judgment, encompassing phonological aspects such as initial and final consonants, vowels, and registers. IPA transcriptions were digitized, and the information was incorporated into an MS Excel table, which included variation and location data with latitude and longitude. The Excel table, now Unicode text (*.txt), was imported into ArcGIS Online, incorporating location details. The data was codified and displayed symbolically and phonological patterns of initial consonants, vowels, and final consonants were considered for mutual variants in each area. Linguistic maps displayed variations, aiding in classifying patterns and analyzing dialect and sub-dialect boundaries.

Results and Discussion

Sub-dialect Classification

This study delved into Khmer dialect geography to identify sub-dialects, focusing on phonological patterns of initial consonants, vowels, and final consonants. These patterns consist of distinct phonological variants used in specific locations. Results reveal a division in the Khmer language spoken in northeastern Thailand into western and eastern forms, particularly concerning vowel variants. Classifying the western and eastern groups relies on vowels like /ə, a, i:, e:, o:/ and /u:ə/. These vowels exhibit clear distribution, with variants distinguishable based on geographical areas in the region. Table 2 presents the variants classifying the vowel patterns.

The data reveals that these vowels distinctly differentiate the western from the eastern regions of the area. Each region has its primary vowel variants, signaling unique patterns. In the western area, variants include 1) i:, 2) e:, 3) o:, 4) a set of u:ə, o:, u:ə, & o:, and 5) a set of ə, o, u. Conversely, the eastern area features 1) i, 2) a set of i:ə, e:, 3) u:, 4) a set of u:ə & ɔ:, u:ə & o:, register vowel, and 5) a. Transition areas share some variants from both the western and eastern regions, leading to the classification of Khmer into the Western group, Eastern group, and Transition areas.

Table 2 The variants for dialect group classification

| Vowel | Variant | | |
|-------|--------------------------------|---|--|
| | Western Group | Eastern Group | Transition area |
| i: | i: | i | uses some variants of Western and Eastern groups |
| e: | e: | 1) i:ə 2) ε: | |
| o: | o: | u: | |
| u:ə | 1) u:ə 2) o: 3) u:ə & o: | 1) u:ə & ɔ: 2) u:ə & o: 3) register vowel | |
| ə | 1) ə 2) o 3) u | - | - |
| a | - | a | - |

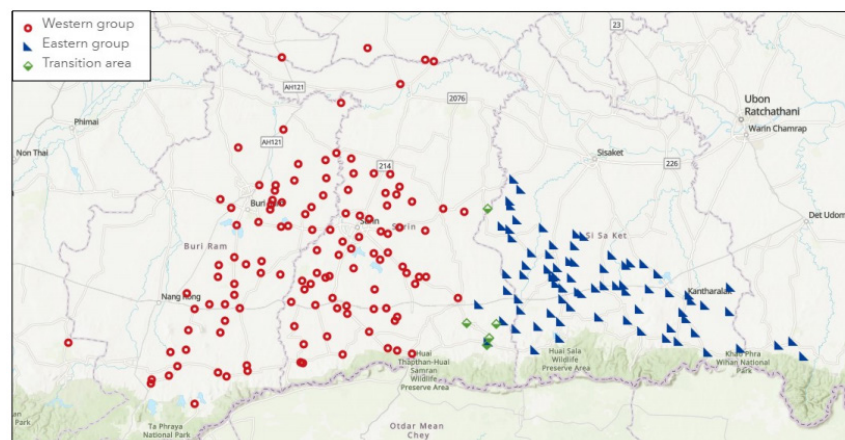
**Figure 1** The main group of Khmer dialects

Figure 1 shows the distribution of dialect groups in the region. The Western group covers the western part, including the areas of Surin, Buri Ram, Roi Et, Maha Sarakham, and Nakhon Ratchasima provinces. The Eastern group covers the eastern part of the region, including the areas of Si Sa Ket and Ubon Ratchathani provinces. The transition area

is the area that uses vowel variants of both groups, but the pattern is not clear enough to be considered a Central group.

The Western and Eastern Khmer groups have main sets of variants that separate them. Within these groups, some phonological patterns according to initial consonant, vowel, and final consonant can help indicate Khmer sub-dialects.

Khmer Sub-dialects According to Initial Consonant

Regarding geographical variation, only four initial consonant phonemes, /c-, n-, ɲ-, and /tc-/, contribute to classifying initial consonant patterns. The patterns consist of a set of distinct variants of initial consonants used in specific locations. The variants that classify these patterns are presented in Table 3.

Table 3 The initial consonant variants for initial consonant pattern classification

| Initial consonant | Variants |
|-------------------|--------------|
| c- | c-, ch- |
| n- | n-, ɲn-, ɲd- |
| ɲ- | ɲɲ- |
| tc- | rc- |

As for the phonological patterns of initial consonants, the results indicate nine patterns. They are:

Pattern 1: c- and n- Pattern 2: c- and ɲn- Pattern 3: c- and ɲd-
 Pattern 4: c- and rc- Pattern 5: ch- and n- Pattern 6: ch- and ɲn-
 Pattern 7: ch- and ɲd- Pattern 8: ch- and rc- Pattern 9: ɲɲ-

Patterns 1, 2, and 3 are commonly spoken in both the Western and Eastern groups. Pattern 4 is mainly spoken in the Eastern group, and only one Western community uses it. Meanwhile, patterns 5, 6, 7, 8, and 9 are only spoken in the Eastern group. The geographical distribution of these initial consonant patterns is shown in Figure 2.

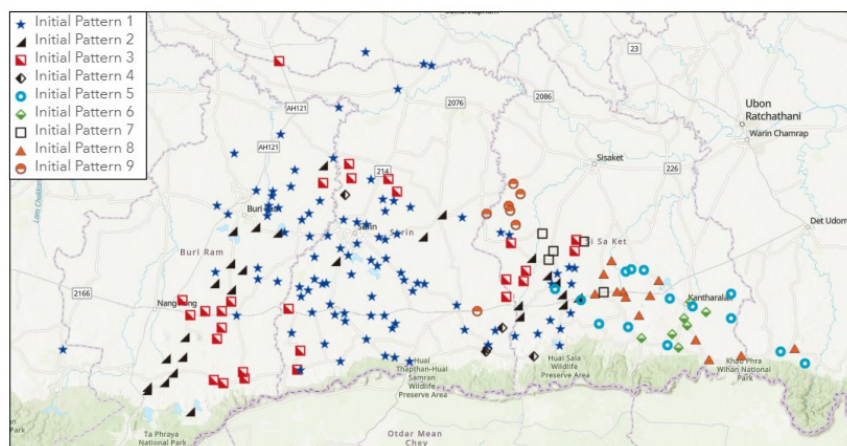


Figure 2 The geographical distribution of initial consonant patterns

Khmer Sub-dialects According to Vowel

According to the phonological pattern of vowel, the Western group has seven patterns of vowel that help indicate seven sub-dialects. The patterns consist of a set of distinct variants of vowels used in specific locations. They are:

- Pattern 1: i - o - e - u:ə - ə Pattern 2: i - o - e - u:ə - o
 Pattern 3: i - o - e - o - ə Pattern 4: i - o - e - o - o
 Pattern 5: i - o - e - o - u Pattern 6: i - o - e - o: / u:ə - ə
 Pattern 7: i - o - e - o: / u:ə - o

The Eastern group has five patterns. They are:

- Pattern 8: i - u - i:ə - u:ə - a Pattern 9: i - u - e - o - a
 Pattern 10: i - u - i:ə - o: / u:ə - a Pattern 11: i - u - i:ə - ɔ: / u:ə - a
 Pattern 12: i - u - i:ə - register

The geographical distribution of vowel patterns is displayed in Figure 3:

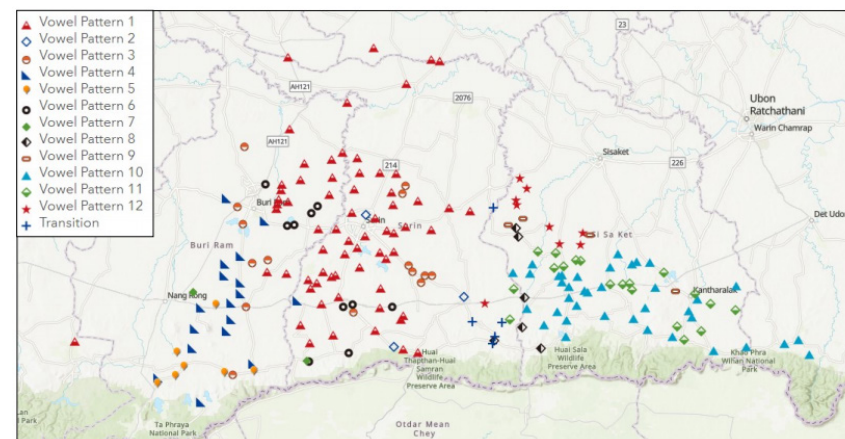


Figure 3 The geographical distribution of vowel patterns

Khmer Sub-dialects According to Final Consonant

The results of the final consonant variation reveal that five final consonants frequently undergo variation: /-c, -k, -ŋ, -r/. The base phoneme, representing the most widespread variant of each final consonant, is presented in the left column, while their respective variants are listed in the right column of Table 4.

Table 4 Final consonant variants

| Final Consonant Phoneme | Variants |
|-------------------------|-------------------|
| -c | -c ~ -k ~ -ʔ |
| -k | -k ~ -ʔ ~ -c ~ -p |
| -ŋ | -ŋ ~ -ɲ |
| -ɲ | -ɲ ~ -ŋ |
| -r | -r ~ -l ~ ɔ |

Concerning pattern classification, the patterns consist of a set of distinct variants of final consonants used in specific locations. It is observed that only the final /-k/ and /-r/ can be effectively utilized. The variants of final /-ŋ/ and /-ɲ/ frequently exhibit variation without

phonological condition. The distribution is mixed, making it challenging to categorize the patterns. Although final /-c/ was employed for classification, it did not contribute significantly to pattern identification. The final -k has four variants: -k, -ʔ, -c and -p. The variants -k and -ʔ are commonly found through the region. In some communities, the variants -k or -ʔ changes to the variants -c or -p according to the phonological condition. The variants -k and -ʔ are found with a single vowel or diphthong. In some communities, these variants changes to the variant -c when they appear after a long front vowel, or vary to the variant -p after a back long vowel. Within these 202 communities, the variant -p is found in only seven communities. The phonological condition of the final -k is:

-k → -k, -ʔ/ all V_ (e.g., *ce:k / ce:ʔ* 'banana')
 -c/ long front V_ (e.g., *ce:k ~ ce:c* 'banana')
 -p/ long back V_ (e.g., *tu:k ~ tu:p* 'boat')

The results indicate the presence of ten patterns for the final consonants. They are:

Pattern 1: -k and -r Pattern 2: -k and Ø Pattern 3: -k and -r/Ø
 Pattern 4: -k and -l Pattern 5: -k/-ʔ and -r Pattern 6: -k/-ʔ and Ø
 Pattern 7: -k/-ʔ and -r/Ø Pattern 8: -ʔ and -r Pattern 9: -ʔ and Ø
 Pattern 10: -ʔ and -r/Ø

Patterns 5 and 7 are spoken in both the Western and Eastern groups. Patterns 2, 6, 8, 9, and 10 are exclusive to the Western group, while patterns 1 and 4 are used in the Eastern group. Pattern 3 is rarely encountered, spoken by only a single community in the Western group and two communities in the Eastern group. The geographical distribution of final consonant patterns is illustrated in Figure 4.

This study affirms that initial consonants, vowels, and final consonants are instrumental in classifying phonological patterns, aiding in identifying dialects and sub-dialects.

Results show that the Khmer language in northeastern Thailand consists of a Western and an Eastern group, primarily distinguished by

vowel patterns. The transition area incorporates shared variants from both groups but lacks an independent pattern. The Eastern and Western groups emerge as the main groups concerning vowels, each containing phonological patterns guiding sub-dialect classification. Sub-dialects within northeastern Thailand amount to 12 for vowels, nine for initial consonants, and 10 for final consonants.

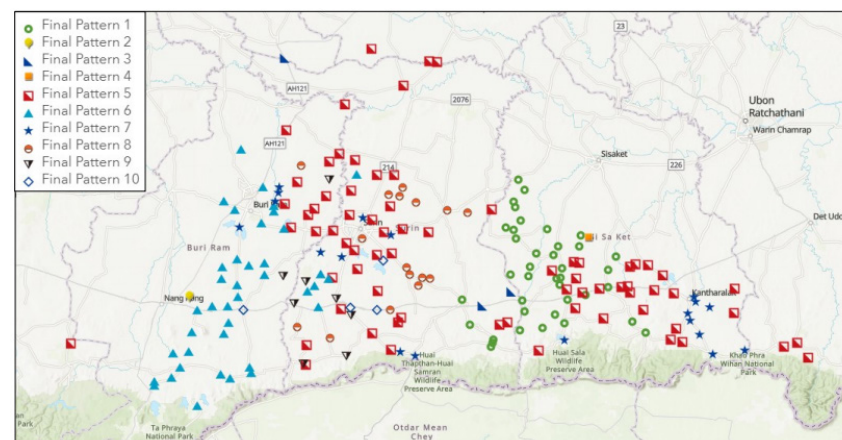


Figure 4 The geographical distribution of final consonant patterns

Khmer Sub-dialects in Northeastern Thailand

The number of sub-dialects in this region varies based on initial consonants, vowels, and final consonants. Vowels provide the largest number of sub-dialects. The distribution of these sub-dialects concerning geographical boundaries indicates proximity to district and province borders but with imperfect fits. Observations reveal a more concentrated sub-dialect distribution at the district level than at the sub-district level. Contiguous districts often share similar features, with sub-dialect boundaries closely aligned with district borders. Some sub-dialects are confined within one district, while others span multiple districts. In contrast, sub-dialect boundaries concerning provincial borders differ. For instance, the Surin dialect extends beyond Surin province into Buri Ram, Roi Et, Maha Sarakham, and Nakhon Ratchasima provinces. It is

suggested to define sub-dialect boundaries as closely aligned with district boundaries. The overview suggests grouping these sub-dialects into three main dialects based on similarities. Therefore, Figure 5 proposes grouping sub-dialects according to district and provincial boundaries.

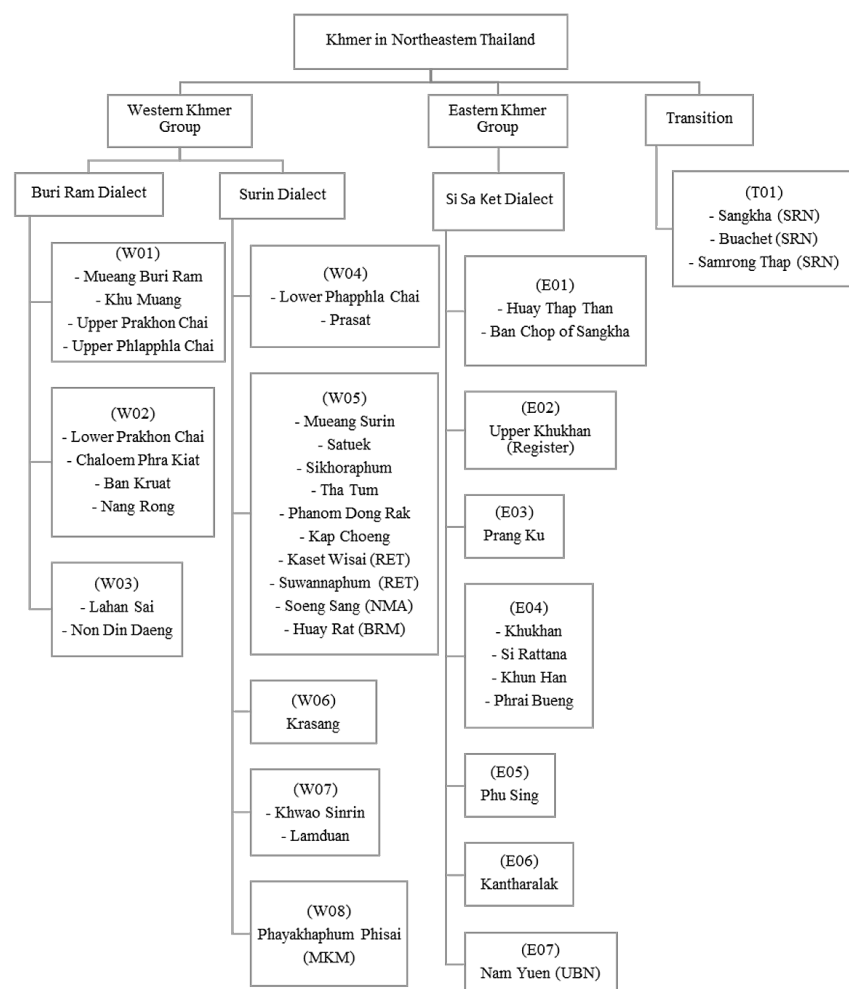


Figure 5 Khmer in Northeastern Thailand²

² The province abbreviations are Ubon Ratchathani (UBN), Si Sa Ket (SSK), Surin (SRN), Buri Ram (BRM), Roi Et (RET), Maha Sarakham (MKM), and Nakhon Ratchasima (NMA).

Figure 5 illustrates the classification of Khmer in northeastern Thailand, delineating the language into two groups. The Western group comprises eight sub-dialects (W01-W08), while the Eastern group encompasses seven sub-dialects (E01-E07). In total, there are 15 sub-dialects within the region. Furthermore, these sub-dialects are consolidated into three main dialects—Surin, Buri Ram, and Si Sa Ket—named after the provinces where they prevail. Each group's delineation is based on the overlap of phonological patterns. Figure 6 depicts the distribution of these fifteen sub-dialects.

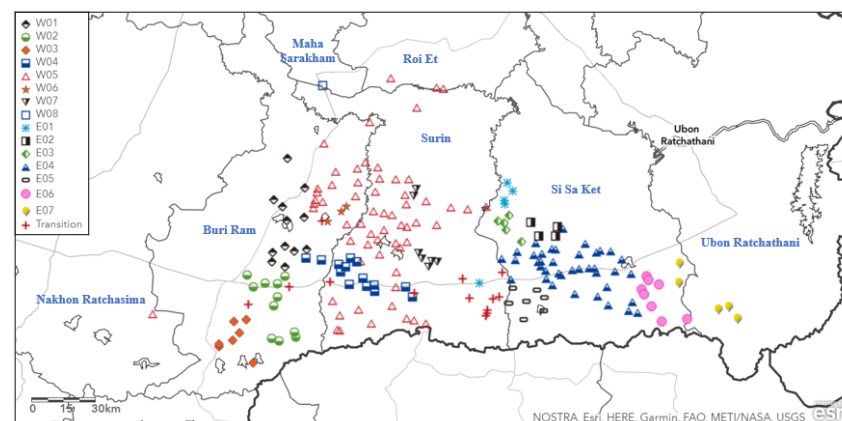


Figure 6 Khmer sub-dialects in Northeastern Thailand

Additionally, the research has identified four key issues related to the causes of language variation, as follows.

Internal Factors

Human languages naturally evolve over time, and it is exceedingly rare for all speakers of a language to use it in exactly the same way. McMahon (1994: 226) stated that language variation is not random, but rather strictly controlled, often by extra-linguistic factors. The specification of these factors may help account for change. Jones and Singh (2005: 1-54) further explained that these linguistic factors, often seen as internal,

drive language change. They stem from the language system itself, such as the desire for easier pronunciation or analogical processes. External factors, on the other hand, come from language contact and include borrowing, convergence, code-switching, and language mixing. In this study, variants of initial and final consonants can indicate variation and sound change due to internal factors. For example:

Example 1

thl- > thl- ~ khl- ~ sl- ~ hl-

These variants indicate the manner-of-articulation changes of plosive *th-*, *kh-* to fricative *s-*, and place-of-articulation changes of the first initial consonants from alveolar *th-* to velar *kh-*.

With the variant *hl-*, an unaspirated liquid consonant, it is assumed that the change is due to the loss of initial consonant *t-* (in *thl-*) or *k-* (in *khl-*) in the first position. So, the *hl-* before vowel is remaining. This is an example of the change that occurs by the language system itself. The phonological conditions indicate the changes of sound that occur due to internal factors.

Sociocultural Factors and Sound Changes

As Labov (1963: 275) emphasized, “One cannot understand the development of a language change apart from the social life of the community in which it occurs.” Although this study primarily focuses on regional variations, the language variation of Khmer inevitably requires addressing social issues.

Sound changes linked to sociocultural factors, such as speaker generations, language contact, migration, and attitudes, indicate potential future language trends. Fieldwork revealed that the younger Khmer speakers pronounce words differently than the older generation. Younger individuals find vowel sounds less complex, using contextual words to convey meaning in sentences instead of relying on sound contrasts. Some registered communities are losing their breathy register, and the final consonant */-r/* has sometimes disappeared.

In Mueang Suea, Maha Sarakham province, the younger generation faces difficulty pronouncing Khmer, leading to a preference

for Thai or the Northeastern Thai dialect. As elderly speakers pass away, fewer people speak Khmer. In some communities, even among older people, the final */-r/* becomes *[-l]*, or */-r/* and */-l/* become *[-n]*. These communities also speak the Northeastern Thai dialect, suggesting that language contact has influenced Khmer changes.

Smalley (1994) described the hierarchy of languages of marginal regional languages in some areas of the lower part of northeastern Thailand, that is: Thai as the official language is at the top in the hierarchy, following by Lao, then other languages such as Khmer and Kuy. Smalley (1994) and Premsrirat (1984) also reported that most of the Khmer speakers can speak at least two languages, Thai and Khmer, and many of them also speak Kuy, Lao, and Thai Korat.

In this study, many communities see the influence of higher-dignity languages, such as Thai—the official language, or the Northeastern Thai dialect—the regional language, both prevalent today. Khmer natives have increasingly adopted Thai, using less Khmer daily, causing older and middle-aged speakers to forget basic vocabulary. The language hierarchy in social prestige descends as follows: Standard Thai (official language), Northeastern Thai or Lao (regional language), other Thai dialects (like Thai Khorat, Thai Nang Rong), Khmer, Kuy, and Ngur. Speakers of lower-prestige languages may also speak higher-prestige languages, but those consistently using high-prestige languages seldom speak lower-prestige ones.

Standard Thai is the primary language for communication in the country, serving as both the national and official language, and is widely spread through mainstream media such as radio and television. As the written language in documents and educational settings, Thai plays a dominant role. Emphasizing the use of Thai in public communication has, at times, led to an overemphasis on ‘Thainess,’ causing local language speakers to prioritize Thai over their native dialects. Languages also differ in social status, reflecting class distinctions. Local dialect speakers often face discrimination, leading to negative attitudes toward their own languages and reluctance to reveal their linguistic identity to avoid social stigma. Therefore, these factors

may contribute to the inability of local languages to assert their presence and significance as they should.

Population mobility is a factor influencing language change. Individuals relocating to new areas often acquire the local languages. In some instances, speakers from another language group, due to marriage, can lead to a shift in the community's language. As the community adopts these introduced languages for daily communication, the usage of Khmer diminishes, resulting in fewer Khmer speakers and alterations in pronunciation or vocabulary.

The speaker's attitude significantly influences language change, impacting the community's language dynamics. In the Sangkha community in Surin province, a migrant group from Si Sa Ket province, the breathy register is preserved across all age groups. Middle-aged speakers, including Khmer, view language as an inheritance, valuing multilingualism. They predominantly speak Khmer at home with their children, while Thai is acquired at school. In contrast, parents in some communities have shifted to speaking primarily Thai at home, believing it aids their children in school, where Thai is the medium of instruction. This shift has led to the younger generation speaking less Khmer.

Geographical factors, like natural barriers such as mountains and rivers, seldom impact language variation, likely due to increased accessibility across the region. With people easily traveling beyond their communities, language interaction extends beyond the geographic constraints of the past. However, observable dialect distinctions arise from distance, evident in the gap between the western and eastern parts of the region.

Language and Migration

The distribution of variants can reveal the connection between language and migration. Most Khmer communities in northeastern Thailand are indigenous, having resided in the region for numerous generations. However, a few communities migrated from other places in the past. For instance, the ancestors of Sangkha Khmer speakers in Surin province moved from Huai Thap Than in Si Sa Ket province around a century

ago. Despite the migration, these two communities speak Khmer with the same sounds, and the register in their vowel system remains preserved. According to members of the Soeng Sang community in Nakhon Ratchasima province, their ancestors migrated from Surin province around 50 years ago, resulting in the Khmer of Soeng Sang being very similar to the Khmer spoken in Surin.

Register and the Khmer Development

The register of the Khmer dialect reveals specific traces, providing insights into its historical development. Changes in Khmer initial consonants contribute to the emergence of the register and diphthongs, highlighting the intricate vowel system. This complexity becomes apparent when the loss of certain vowels results in voicing changes in initial voiced stop consonants.

Historically, Khmer featured a register to distinguish word meanings (register contrast), categorizing consonants based on voicing. Phonological changes arose due to the gradual loss of voicing in initial consonants. When an initial consonant loses its voicing contrast, the register shifts to vowels, distinguishing word meanings through a breathy and clear vowel system. Later, the vowel loses its register contrast, leading to the merging of vowel groups. Consequently, one of the vowel groups must alter its sound to maintain phonological contrast (vowel contrast). This complexity in the vowel system arises to compensate for the missing register contrast (Jenner, 1974; Pisitpanporn, 2005), resulting in Khmer's extensive vowel inventory.

In northeastern Thailand, Southong (2013) noted a community in Si Sa Ket province speaking Khmer with a breathy register. This study validates these findings, extending the scope by collecting sub-district-level data from numerous areas. The register persists in nine communities in Si Sa Ket province, including four in Huai Thap Than, four in Khukhan, as well as one in Sangkha, Surin province. The Sangkha community, having migrated from Si Sa Ket around a century ago, still preserves the register in vowels.

The ongoing process of transitioning from register contrast to vowel contrast is assumed. A comparison of registered vowels across the nine communities revealed that only speakers in Chan Saen Chai consistently used breathy vowels for every word on the list. In other communities, speakers employed breathy vowels, clear vowels, or diphthongs for various words, even when the register was present.

Conclusion

Variations within Khmer occur in all phoneme aspects: initial consonants, final consonants, and vowels. Sets of variants play a pivotal role in classifying phonological patterns, aiding in identifying dialects and sub-dialects. Register, integral to vowels, emerges as a significant feature in nine communities—eight in Si Sa Ket province and one in Surin province—where Khmer is spoken with a breathy register in their vowel system. Final phoneme variation, notably in five Khmer consonants, shows predictability and clarity compared to initial consonants and vowels, often presenting two to four variants.

Among all consonant and vowel phonemes, only three final consonants—/c/, -k/, and /-r/—exhibit variations under phonological conditions. Some variants are widespread across the region, while others are specific to either the western or eastern areas. For instance, the vowel variant [e] is prevalent throughout the region, while [ɛ] is exclusive to the east part. Consequently, the presence of this variant in a particular area serves as a classification marker for the dialect group.

The Khmer dialects can be categorized into two primary groups—Western and Eastern—based on vowel patterns. Sub-dialects are distinguishable by analyzing phonological aspects, including initials, vowels, and finals. The Western Khmer group exhibits seven vowel patterns, while the Eastern Khmer group has five. The Transition area shares mutual variants from both groups but lacks distinct patterns. Consequently, the Eastern and Western groups emerge as the main areas concerning vowels.

Sub-dialect classification, incorporating all aspects (initial consonants, vowels, and final consonants), demonstrates noteworthy similarities. The distribution overview indicates comparable sub-dialect groupings, with specific patterns identified as sub-dialects. Thus, Khmer dialects in northeastern Thailand are known to encompass twelve sub-dialects for vowels, nine for initial consonants, and ten for final consonants.

The distribution maps highlight distinctions in Khmer dialects between the western and eastern regions of the area. The gap between the borders of Si Sa Ket and Surin provinces signifies the separation between Western Khmer and Eastern Khmer dialects. Although political and dialect boundaries do not align perfectly in this region, district or provincial borders can serve as language boundaries in some areas. Thus, it may be more fitting to name sub-dialects according to district names.

References

- Chambers, J. K. and Trudgill, P. (1980). **Dialectology**. Cambridge: Cambridge University Press.
- Chantrupanth, D. and Phromjakgarin, C. (1978). **Surin Khmer-English dictionary**. Bangkok: Institute of Indigenous Language Research, Chulalongkorn University Press.
- Crystal, D. (1992). **An encyclopedic dictionary of language and languages**. Oxford, UK: Blackwell.
- Hattaway, P. (Ed.). (2004). Khmer. In **People of the Buddhist world: A Christian Prayer Diary**. (pp. 133). Pasadena, CA: William Carey Library.
- Jenner, P. N. (1974). The development of the registers in standard Khmer. In N. D. Liem (Ed.). **Vietnamese and other Mon-Khmer languages**. (pp. 47-60). Canberra: ANU Asia-Pacific Linguistics / Pacific Linguistics Press.
- Jones, M. C. and Singh, I. (2005). **Exploring language change**. London: Routledge.
- Kraisornrat, S. (2012). *Watthanatham paet chattiphan changwat Ratchaburi nai boribot khong kan hhongthiao*. (In Thai) [The culture of 8 ethnographic groups in Ratchaburi province in the context of tourism]. **Journal of International and Thai Tourism**, 8(2), 19-31.
- Labov, W. (1963). The social motivation of a sound change. **Word**, 19, 273-303.
- McMahon, A. (1994). **Understanding language change**. Cambridge: Cambridge University Press.

- Phon-ngam, P. (1987). **A phonological comparison of spoken central Khmer (Phnom Penh) and northern Khmer (Surin)**. M.A. Thesis, Mahidol University.
- Pisitpanporn, N. (2005). *Kanphatthana wannayuk nai phasa Khamen Phanom Pen*. (In Thai) [Tone Development in Phnom Penh Khmer]. **Journal of Mekong Societies**, 1(3), 107-120.
- Premssirat, S. (1984). Culture in language use (The use of address terms in Northern Khmer). **Journal of Language and Culture**, 4(1), 24-29.
- Premssirat, S. (1997). Linguistic contributions to the study of the northern Khmer language of Thailand in the last two decades. **Mon-Khmer Studies**, 27, 129-136.
- Premssirat, S. and Phuitharakit, O. (1996). **Laksana lae kan krachai khong phasa khamen thin Thai adit patchuban anakhot khong phayanchana sakot**. (In Thai) [The characteristics and distributions of northern Khmer: The past, present, and future of some dynamic final consonants]. Nakhon Pathom: Research Institute for Languages and Cultures for Rural Development, Mahidol University.
- Premssirat, S. and Srichampa, S. (1990). **Kan phatthana rabop kan khian phasa khamen thin Thai**. (In Thai) [Northern Khmer writing system development]. Nakhon Pathom: Institute of Language and Culture for Rural Development, Mahidol University.
- Premssirat, S., Diphadung, S., Suwanket, E., Chusi, I., Sichampa, S., Buasuang, A., Thawon phat, M., Thawisak, A., and Damsa-at, P. (2004). **Phaenthi phasa khong klum chattiphan tang tang nai prathet Thai**. (In Thai) [Ethnolinguistic Maps of Thailand]. Bangkok: Office of the National Culture Commission.
- Seedabut, P. (2016). **Rabop siang phasa khamen thin Thai tambon Ta Khian Ram amphoe Phu Sing changwat Si Sa Ket**. (In Thai) [A phonological study of a Khmer dialect at Takhian Ram sub-district, Phu Sing district, Si Sa Ket province]. M.A. Thesis in Linguistics, Mahidol University, Thailand.
- Seedabut, P. and Burusphat, S. (2017). *Kan priapthiap rabop siang phasa khamen thin thai changwat Si Sa Ket kap thin Surin Buri Ram Chanthaburi lae Chachoengsao*. (In Thai) [A phonological comparison of Si Sa Ket Khmer dialect with Surin, Buri Ram, Chanthaburi, and Chachoengsao dialects]. **Journal of Human Sciences**, 18(1), 12-39.
- Smalley, W. A. (1994). **Linguistic diversity and national unity**. Chicago: The University of Chicago Press.
- Southong, C. (2002). **Kansueksa rabop siang phasa Khamen ban Samrong Bon tambon Saikhao amphoe Soidao changwat Chanthaburi**. (In Thai) The phonological study of Khmer spoken at ban Samrongbon, tambon Saikhaw, amphoe Soidao, Chanthaburi province. M.A. Thesis, Silpakorn University, Thailand.
- Southong, C. (2013). **Vowel variation of Khmer in Thailand**. Unpublished Ph.D. Dissertation, Mahidol University, Thailand.
- Thomas, D. (1992). On sesquisyllabic structure. **Mon-Khmer Studies**, 21, 206-210.
- Treesawat, W. (2012). *Wat Mon changwat Ratchaburi: Kan chat Phumithat lae khwammai khong attalak chumchon*. (In Thai) [Mon temples in Ratchaburi province: Landscape designs and definitions of community identity]. **Romphruek Journal Krirk University**, 30(2), 1-28.
- Vail, P. (2007). Thailand's Khmer as 'invisible minority': Language, ethnicity and cultural politics in north-eastern Thailand. **Asian Ethnicity**, 8(2), 111-130.
- Wichitkhachee, A. (1996). **A phonological study of the Khmer spoken at Ban Nawattai, tambon Dongnoi, Ratchasan sub-district, Chachoengsao province**. M.A. Thesis, Mahidol University, Thailand.
- Yodmongkhon, K. (1986). **The phonological study of northern Khmer Prakhonchai district, Buriram province**. M.A. Thesis, Mahidol University, Thailand.

Websites

- Cultural Center of Mueang Nakhon Pathom. (2013). **Ethnicity in Nakhon Pathom province**. Retrieved October 15, 2018, from http://service.christian.ac.th/ncc/Ethnic/Kamain/Kamain_Index.html
- Esri. (n.d.). **ArcGIS Online**. Retrieved September 5, 2024, from <https://www.esri.com/en-us/arcgis/products/arcgis-online/overview>
- Suphanburi National Archives. (2018). **Moon festival at Samthong temple, Taling Chan, Mueang Suphanburi district, Suphanburi province**. Retrieved October 15, 2018, from <https://www.finearts.go.th/promotion/view/25309>

Appendix

Summarized details of each dialect's phonemic systems and phoneme inventories are as follows:

Phonemes of Khmer dialects in Thailand

| Dialect | Initial consonant | Final consonant | Vowel |
|--|--|--|--|
| Surin (Chantrupanth and Phromjakgarin, 1978) | p t c k ʔ b d ph th ch kh m n ɲ ŋ r l w j s h (21 phonemes) | p t c ʔ m n ɲ ŋ r l h (11 phonemes) | i i: u u: u u: ɪ ɪ: ʌ ʌ: w w: e e: ə ə: o o: ɛ ɛ: ʌ ʌ: ɔ ɔ: a a: ɑ ɑ: ei ei ai a:i ɛ:i ɔ:i ʌi ʌ:i ui w:i oi ɔ:i ɑ:i iu i:u eu e:u ɛ:u au a:u ɯ u ʌ u ia i:a ua u:a ua u:a iai uai iau (60 phonemes) |
| Surin (Phon-ngam, 1987) | p t c k ʔ b d m n ɲ ŋ r l w j s h (17 phonemes) | p t c k ʔ m n ɲ ŋ r l w j h (14 phonemes) | i i: u u: u u: ɪ ɪ: ʌ ʌ: w w: e e: ə ə: o o: ɛ ɛ: ʌ ʌ: ɔ ɔ: a a: ɑ ɑ: ia i:a ua u a u:a (33 phonemes) |
| Surin (Premsrirat and Srichampa, 1990) | p t c k ʔ b d ph th ch kh m n ɲ ŋ r l w j s h f (22 phonemes) | p t c k ʔ m n ɲ ŋ r l w j h (14 phonemes) | i i: u u: u u: ɪ ɪ: ʌ ʌ: ɔ ɔ: e e: ə ə: o o: ɛ ɛ: ʌ ʌ: ɔ ɔ: a a: ɑ ɑ: ia i:a ua u:a ua u:a (34 phonemes) |

| Dialect | Initial consonant | Final consonant | Vowel |
|--|--|--|---|
| Buri Ram (Yodmongkhon, 1986) | p t c k ʔ b d ph th ch kh m n ɲ ŋ r l w j s h (21 phonemes) | p t c k ʔ m n ɲ ŋ r l w j h (14 phonemes) | i i: u u: u u: ɪ ɪ: ʌ ʌ: w w: e e: ə ə: o o: ɛ ɛ: ʌ ʌ: ɔ ɔ: a a: ɑ ɑ: iə i:ə uə uə (31 phonemes) |
| Chanthaburi (Southong, 2002) | p t c k ʔ b d m n ɲ ŋ r l w j s h (17 phonemes) | p t c k ʔ m n ɲ ŋ r l w j h (14 phonemes) | i i: u u: u u: e e: ə ə: o o: ʌ: æ æ: a a: ɔ ɔ: iə uə uə əu i i: ɯ ɯ: ɯ ɯ: e e: ə ɔ ɔ: æ: ʌ ɔ ɔ: iə ɯə ɯə (41 phonemes) |
| Chachoengsao (Wichitkhachee, 1996) | p t c k ʔ b d m n ɲ ŋ r l w j s h (17 phonemes) | p t c k ʔ m n ɲ ŋ l w j h (13 phonemes) | i i: u u: u u: ɪ ɪ: ʌ ʌ: w w: e e: ə ə: o o: ɛ ɛ: ʌ ɔ ɔ: a a: ɑ ɑ: ia i:a ua u:a ua u:a a:ə a:ə a:ə (35 phonemes) |
| Si Sa Ket (Seedabut, 2016) | p t c k ʔ b d m n ɲ ŋ r l w j s h (17 phonemes) | p t c k m n ɲ ŋ r l w j h (13 phonemes) | i i: u u: u u: ɪ ɪ: ɔ ɔ: e e: ə ə: o o: ɛ ɛ: ʌ ʌ: ɔ ɔ: a a: iə i:ə uə uə (28 phonemes) |